A Safety Index and Method for Flightdeck Evaluation.

Kara A. Latorella
Crew/Vehicle Integration Branch, NASA Langley Research Center

ABSTRACT
If our goal is to improve safety through machine, interface, and training design, then we must define a metric of flightdeck safety that is usable in the design process. Current measures associated with our notions of "good" pilot performance and ultimate safety of flightdeck performance fail to provide an adequate index of safe flightdeck performance for design evaluation purposes. The goal of this research effort is to devise a safety index and method that allows us to evaluate flightdeck performance holistically and in a naturalistic experiment. This paper uses Reason's model of accident causation (1990) as a basis for measuring safety, and proposes a relational database system and method for 1) defining a safety index of flightdeck performance, and 2) evaluating the "safety" afforded by flightdeck performance for the purpose of design iteration. Methodological considerations, limitations, and benefits are discussed as well as extensions to this work.

INTRODUCTION

This research is motivated by the question “How do we know if we are designing a safer flightdeck?” This effort aims to measure the degree to which a flightdeck supports good pilot performance. It is assumed that by supporting good pilot performance the goals of pilot performance, principally “safety,” are better achieved. This section reviews current approaches to measuring aviation safety and identifies requirements for a *flightdeck performance safety index*. Based on these considerations, I present a candidate safety index and method for evaluating flightdeck performance safety.

Measuring Safety

If our goal is to improve safety through machine, interface, and training design, then we must define a metric of flightdeck safety that is usable in the design process. Accident rate, in terms of hull losses or fatalities, has been used as a measure of aviation system safety. However, accidents are extremely rare occurrences, and therefore are not a meaningful metric for evaluating the safety of a particular flight/flightdeck/operator mission. Incidents, including regulatory violations, while more frequent than accidents are still relatively rare (Wickens, 1995, p. 126). Typical human performance experiments in aviation assess the degree to which a particular system, procedure, or training regime improves safety by measuring a narrow band of performance measures directly related to that intervention. These focused experiments typically include system-specific errors and a few particular reaction times as dependent measures. There are several problems with using this approach. First, subjects manage trade-offs between reaction-time performance and accuracy (Pachella, 1974). Second, these finer grained measures can be of questionable operational significance. For example, faster reaction times do not always translate to behavior consistent with improved safety (*cf*. Rogers, Schutte and Latorella, 1996). Finally, in these focused experiments, you frequently get what you measure; that is, subjects manage resources to optimize performance on those aspects that they believe are being measured. Subjective assessments are also commonly used in focused experiments. However, subjective evaluations may dissociate from actual performance (Yeh & Wickens, 1988) and may be unfairly biased towards familiar designs. Physiological measures of stress and arousal only directly indicate intermediary states, and are therefore also removed from an operational definition of safe performance. In summary, current measures associated with our notions of "good" pilot performance and ultimate safety of flightdeck performance fail to provide an adequate index of safe flightdeck performance for design evaluation purposes.

Requirements for a Safety Index

The goal of this research effort is to devise a safety index and method that allows us to evaluate flightdeck performance holistically and in a naturalistic experiment (Beach *et al.*, 1997). This safety index must provide sufficient data to indicate levels of safety achieved in flightdeck performance. The index must be *traceable* to conventional notions of safety in aviation to support *face and construct validity*. The safety index should be *sensitive*; that is, it should provide an adequate range of magnitude to reflect differences in what is measured. It should be *unobtrusive* and *non-reactive*; that is, it should not interfere with the natural performance of the task, and it should not be obvious to the subject what is being measured. By ensuring that the measure is not obtrusive or invasive and does not induce biases to "game" performance, one encourages subjects' motivational and goal structures in the real environment to be retained in the experimental environment. Finally, the index must be *generalizable*; that is, usable across classes of flightdecks for comparison purposes, and tailored to a particular mission. The safety index, to be a worthy measure, must also posess *predictive validity* and be *reliable*.

A SAFETY INDEX

Theoretical Underpinnings

To define this safety index, I revisited Reason's model of accident causation (1990). Reason's model depicts accident causation as the result of latent and active failures penetrating barriers erected to prevent their culmination in an accident. This model describes what has been demonstrated in many accident reports: accidents are not usually caused by one isolated event. Accidents emerge from the combination of smaller errors, smaller hazards, subtle precursors, and latent weaknesses. To simplify the aforementioned model for a specific point, one could say there is a semi-permeable membrane, an imperfect final defense, between "unsafe-acts" (Reason, 1990) and accidents. It seems reasonable to postulate, then, that the concentration of unsafe-acts on the forcing side of that membrane should be predictive of the likelihood of a transfusion through it, and therefore the probability of a resulting accident. Hollnagel (1991) has made a similar observation, stating that as the complexity of an interface increases, the number of erroneous actions will increase, rather than the types of erroneous actions. I define a potential unsafe-act (PUA) as a deviation from prescribed flightdeck performance based on a competency model of piloting. I assume that pilots' intentions are to act in accordance with this model.

Defining Terms in the Safety Index

For any evaluation, only some PUAs from the set of all possible PUAs will be applicable to the mission scenario. This subset of PUAs is partitioned, based on pilot performance, pilot competence and pilot intent, into quantitative terms used in the safety index. These terms are defined below.

Correct Actions (CA): First, those PUAs that a pilot does not commit during an evaluation are considered Correct Behaviors (CB) and indicate behavior consistent with the prescribed performance model. CAs are those correct behaviors for which the pilot *knew* the required performance and acted accordingly.

Expressed and Unexpressed Competency Errors (CE): Strictly using a set of PUAs to evaluate performance assumes that pilots have knowledge of the model from which these PUAs are derived. However, we require a safety index that reflects performance decrements associated with the ability of the flightdeck design to support operator performance, rather than deficits in pilots' knowledge. A safety index of a flightdeck ought to be independent of the knowledge inherent in the operator - although the ultimate assessment of "safe" performance will necessarily be a function of the two. Therefore performance deviations must be distinguished from competency errors (*cf.* Chompsky, 1957; Smith and Hancock, 1995; Joseph and Uhlarik, 1997). Competency Errors (CE) are the subset of PUAs for which the pilot did not know the correct behavior, and is considered incompetent. CEs may or may not result in a performance deviation. Where a CE does not emerge as a performance deviation, it is considered *unexpressed* (CEU) and is a subset of correct behaviors. *Expressed* competency errors (CEE) are a subset of committed PUAs, *i.e.*, performance deviations (PD).

Innovative Acts (IA): Simply using this residual set of performance deviations as an index of flightdeck safety assumes that all performance deviations indicate inappropriate behavior. In the aviation domain one can be reasonably sure that subjects will not intentionally perform violations. However, other circumstances may foster intentional deviations from prescribed performance. According to a strict interpretation, adaptive behaviors to unusual situations, and acceptable alternative methods would be considered intentional performance deviations - violations. Underspecificity and incompleteness of the competency model and derived set of PUAs would result in spurious intentional performance deviations. Definition of IAs, therefore, is also important for methodological purposes. In this formulation, performance deviations are considered IAs only if they are successful innovations or acceptable alternative methods. Innovative Acts are isolated as a second subset of PDs.

Operational Errors (OE): After removing all CEE and IA from the set of PDs, only OEs remain. Operational Errors, then, are defined as performance deviations from a prescribed model of pilot competence that are neither attributable to competency errors nor successful innovation.

Formulation

This safety index evaluates the degree to which a flightdeck performance supports prescribed behavior. An additional assumption of the proposed index is that it is an advantage of a flightdeck to support successful adaptive behavior and use of acceptable alternative methods. Finally, the safety index assumes that flightdeck designs that compensate for pilot competency errors are safer. Therefore, the safety index (SI) for a scenario (s) and a particular subject (p) is defined by the terms above as:

$$SI_{(s,p)} = (CA_{s,p} + IA_{s,p} + CEU_{s,p}) / PUA_s$$

The stability of a flighdeck evaluation with SI improves by evaluating $SI_{(s,p)}$ over a large number of subjects (P) with a breadth of individual subject characteristics (*e.g.*, experience, personality traits, *etc.*), and a variety of mission scenario elements (*e.g.*, weather phenomena, terrain, system events). The averaged $SI_{\bullet\bullet}$, then is a general index of safety for flightdeck performance. Note that the formula for this safety index is undefined when CEEs equal prescribed performance for PUAs. This is not a likely occurrence in practice (and if it is, we have bigger problems than the flightdeck design), however it is a limitation of this formulation. The remainder of this paper describes a database system and an evaluation method that support the use of this safety index for evaluating flightdeck performance.

PUAs & OTHER DATABASES

This section describes development of the Potential Unsafe-Acts (PUAs)[1] database as well as, briefly, other databases required to support use of the safety index.

The PUA Database

A competency model of performance is analyzed for potentially unsafe acts. This section describes the process of identifying PUAs and the information associated with each PUA in the database. Table 1 presents an example PUA database item. The fields associated with database items are described below in order of their appearance in Table 1.

Table 1. Example PUA Database Item

AIM Section: 5.2.6.b.3
Knowledge Element: Climb gradients are specified when required for obstacle clearance. Crossing restrictions in the SID's may be established for traffic separation or obstacle clearance. When no gradient is specified, the pilot is expected to climb at least 200 feet per nautical mile to MEA unless required to level off by a crossing restriction.
Behavioral Rule: If {(no gradient for obstacle clearance is specified) and (not required to level off by a crossing restriction)} Then (climb at a rate of at least 200 ft/nm to the MEA).
Context Factors: aircraft type (All), visibility (IFR), airspace class (Any), airspace type (General), flight phase (In-Flight), weather (none), crew (acceptable), emergency (irrelevant), hazard (no), operating conditions (normal), day/night (irrelevant) <u>Antecedent Factors:</u> - no gradient specified for obstacle clearance - no requirement for level off at a crossing restriction
PUA(s): 1. Inaccurate performance: climbing < 200'/nm to MEA
Observation Requirements/Scenario Flags: <u>Context Appropriateness Requirements:</u> IFR, In-flight <u>Error Observation Requirements:</u> Climb rate
Competence Assessment Item(s): 1. If no gradient is specified for obstacle clearance, and you aren't required to level off at a crossing restriction, what is your minimum climb rate to the MEA?

A Competency Model of Pilot Performance: Numerous requirements exist for the competency model source. It should cover the breadth of a generic piloting mission. It should avoid company or aircraft-specific performance objectives, such as defined in standard operating procedures, to ensure generalizability across populations. These two goals suggest a source that describes performance at a relatively high level. Finally, the source must describe desired performance to enable expression of observable performance errors and contexts in which they occur. A true competency model of piloting would include aspects of psychomotor skill, rules for performance, as well as declarative knowledge of aerodynamics, the airspace, and regulations. The first iteration in this effort, focuses on codifying rule-based behavior required for successful performance; the information determined by the constraints of declarative knowledge and that guides skilled performance.

After considering several FAA (Federal Aviation Administration) sources (*e.g., Federal Aviation Regulations* (FARs), *Airman Practical Test Standards*, *Knowledge Tests*), it was determined that the *Aeronautical Information Manual* (AIM) best met these objectives. The AIM provides an overview description of good piloting behavior that is neither aircraft model, nor company specific. The AIM provides the aviation community with basic flight information and ATC procedures for use in the US National Airspace System (NAS). While intended to be supplemented by more current information in *NOTAMS* (notices to airmen) and the *Airport/Facility Directory*, and not intended to be regulatory, the AIM provides operating techniques and procedures that indicate appropriate and legal piloting behavior. The AIM was selected as the initial source document for this research effort for its generality, mission coverage, and focus on appropriate performance methods. A limitation of using this reference is that it only addresses standard conditions. Unusual circumstances may require pilots to deviate from these standard behaviors. While other sources were deemed less appropriate for the initial objectives of this study, they all contain useful information that should be added as this database evolves, with review by expert subjects.

Knowledge Elements & Behavioral Rules: The entire text content of the AIM (v. 2/26/98) was canvassed to select knowledge elements. The associated pilot/controller glossary was not reviewed. Information presented graphically or in tabular format was not included unless not represented by the text. Knowledge elements in the AIM were defined, for our purposes, as statements that prescribed pilot performance. These usually were stated as a recommendation for the pilot and were surrounded by the contextual conditions for which this behavior was prescribed. The AIM was reviewed for these knowledge elements by considering whether not performing the prescribed action, or performing it incorrectly could conceivably be construed as a contributing factor in an accident investigation. This broad definition resulted in the initial definition of 1629 knowledge elements. These knowledge elements are currently under review.

[1] The PUA database is implemented by NASA contract NAS1-96014, DC22R2 with Lockheed-Martin (see Press, 1998).

Behavioral rules are defined for each knowledge element. The antecedents for each rule specify a characteristic of the simulation scenario. The consequent of the rule specifies a performance goal for this scenario. The conditionals, as stated in this paragraph are not sufficient for describing the context under which this rule is appropriate. For this reason, other context factors were defined to identify broader situations in which a rule applies.

Context Factors: The knowledge elements were described by the broader context in which the prescribed behavior is appropriate. These factors were not intended to represent exclusive categories. Rather, these terms were selected to describe conditionals expressed in the AIM, and to facilitate experimenters in building scenario conditions. The eleven context factors are: aircraft-type, visibility-condition, airspace-class, flight-hazard, equipment-usage, emergency-condition, weather-types, crew-condition, special-airspace-types, phase-of-flight.

Potential Unsafe Acts (PUAs): Simulation evaluations are constrained to detecting observable errors, therefore only error phenotypes are identified here. Each knowledge element will be evaluated for PUAs using Rasmussen *et al.*'s (1981) classification for phenotypes associated with the "external mode of malfunction" (p. 93) (Table 2). The expression of PUAs is limited to simple phenotypes because these are accumulated in the method.

Table 2. Simple Phenotypes (Rasmussen, 1991).

Specified task not performed - Omission of act - Inaccurate performance - Wrong timing (premature action,delay)
Commission of erroneous act(repetition,reversal,replacement)
Commission of extraneous act (intrusion)

Data Collection and Sensing Requirements: To use PUAs in simulation evaluation, one must identify how to detect when they are appropriate ("context-appropriateness requirements"), and when they are committed ("observation requirements"). These may be either flags set *a priori* describing the scenario or may be defined in terms that can be measured dynamically during testing.

Competency Assessment Items: To address this issue, our method requires that, following simulation testing, subjects be assessed for their competency. "competency assessment items" for each PUA to ascertain whether subjects have the requisite knowledge for competent flightdeck performance.

<u>*Other Databases*</u>

In addition to the PUA database, this method assumes the existence of several other linked databases. The Subject Database has a frame for each individual participant containing demographic, scores, and test participation information. The Simulation Database contains a mapping of PUA Database sensing requirements to simulation variables, and routines and links to the missions defined in the Evaluation Database. The Evaluation Database has a frame for each evaluation containing test identification information, mission description, summary scoring, IAs, and links to participants in the Subject Database.

EVALUATION WITH THE SAFETY INDEX

This section focuses on how designers interact with the PUA database to evaluate flighdeck performance with the safety index.

<u>*Step 1: Designing the Evaluation*</u>

First, the researcher describes the mission scenario using the context factors and antecedent terms. The database selects the PUAs set defined for this mission.

<u>*Step 2: Designing the Simulation*</u>

For each of the potential unsafe-acts, the database also provides the researcher with data requirements for sensing the occurrence of these acts. Some of these data requirements are as simple as detecting a button push. Others will require calculations of time elapsed, distance traversed, *etc.* Still others will require additional sensing capability. One example of this might be a requirement for an eye-tracking device to determine data requirements such as "pilot detects." Clearly the process of providing operational definitions for data requirements is non-trivial. However there are clear benefits to maintaining a consistent database of these definitions to facilitate comparison across studies. Researchers will be allowed to eliminate certain PUAs from the evaluation set if inclusion of data requirements is prohibitive. Through use with a specific flightdeck simulation program, data requirements will, over time, be associated with simulation variables and routines, reducing the programming effort required for data collection. At this step, the PUAs for this mission are established and defined in the simulation software for data collection.

<u>*Step 3: Data Collection*</u>

During experimentation, the simulation software executes the rules for defining operational errors and posts these time-stamped PUAs when they occur. At the conclusion of an evaluation, the sets of performance deviations (PDs) and correct behaviors (CBs) are defined for the subject. In addition to the simulation data collection, video recording of subjects' performance is recommended to inform IA identification.

<u>*Step 4: Identifying Competency Errors*</u>

Based on the initial set of PUAs, the researcher uses the database to format competency assessment items into a scenario-specific questionnaire. After evaluation, a subject completes this questionnaire to identify those scenario PUAs for which he did not know the correct behavior. Results of this questionnaire in conjunction with performance define CEEs and CEUs.

Step 5: Operational Errors v. Innovative Acts
As part of the evaluation debriefing, the researcher and subject review performance deviations not due to CEE to identify IAs. IAs are recorded and linked to subject and scenario descriptions. Remaining performance deviations are considered OEs for this mission scenario and subject's flightdeck performance.

Step 6: Assess Safety Index
To summarize, the original PUA database is tailored by the researcher to a smaller set specific for a particular evaluation scenario. After evaluation, this subset of PUAs is partitioned (CA, CEE, CEU, IA, OE) and a SI is calculated for a subject/scenario combination. These elemental SIs are analyzed over many subjects and scenarios.

DISCUSSION

This paper proposes a metric and a companion methodology for evaluating the safety of flightdeck performance. Although this effort is still underway, some methodological considerations, benefits, and extensions of this approach are apparent.

Methodological Considerations
This approach relies on a competency model of piloting. Although the AIM is a useful starting point for this effort, it has some limitations. First, it generally addresses rule-based knowledge for event-driven behavior. It does not detail the level of psychomotor skill required to execute maneuvers, nor does it contain the wealth of declarative knowledge that pilots use to tailor standard practices. So, potential errors such as a deviation from the flightpath of some significant magnitude is conspicuously absent from this PUA database. Similarly, evaluation using this database requires researcher expertise to identify successfully adaptive behavior. Secondly, the AIM addresses generic pilotage and fundamentally pertains to behaviors supporting the goals of "aviate, navigate, communicate." However, pilot<u>ing</u> also requires behaviors in support of other goals (task management, mission management, passenger management, *etc.*). While errors of these forms are not immediately associated with safety, they could contribute to sub-optimal flightdeck performance. Finally, the AIM is expressly for current technology and conventions. It is appropriate for the current National Airspace System (NAS) infrastructure, operating regulations, and equipment and addresses use of specific equipment onboard current aircraft. Therefore using only the AIM as a competency model source limits extendibility of this method to new systems or airspace environments.

A second methodological consideration addresses the derivation of the safety index. The definition of the safety index is only an index of observable "unsafe-acts." This positivist stance requires that a problematic condition express itself outright before it is considered problematic. Conditions that are traditionally considered problematic in and of themselves because they have been established as precursors to negative observable consequences are not counted. Consider the issue of workload. The safety index does not address workload directly. Rather, it assumes that if workload is high enough to present a problem, this problem will become observable as an operational error and counted in the safety index. By not explicitly considering workload, we do not understand the level of effort required on the human operator's part to achieve the observed level of system performance. This is of critical import if this safety index is to be used for flightdeck comparison evaluations and is therefore a limitation of this method.

Another concern relates to the implementation of the safety index. Designers, including the designers of this database, are not above committing human error. Errors in identifying knowledge elements, defining the logical requirements for rules, identifying sensor and data requirements, and writing unbiased competency items are potential failure points in this system. Finally, our most pragmatic concern is whether the data obtained from using this safety index achieves the *a priori* requirements of sufficiency and sensitivity. Will there be enough unsafe-acts to count? Will there be any remaining operational errors after removing IA or competency errors? Will this index be robust enough to overpower individual difference effects? Such concerns will only be addressed by exercising this index empirically in simulation experiments.

Assessments using this safety index are limited by how fully the evaluation scenarios exercise performance-forcing conditions resident in the real operational domain. To fully exercise these conditions, one would evaluate a flightdeck interminably. Obviously this is prohibited by resource constraints and need for <u>an</u> assessment, if not a perfect assessment.

Benefits
Even in light of these limitations, the benefits of this approach are numerous. It provides a naturalistic, quantitative means for assessing the safety of a human-machine system based on operationally-relevant behaviors. This method facilitates the process of conducting a simulation evaluation by explicating the relationship between scenario design and evaluation requirements, and by specifying requisite data and sensing requirements. It facilitates comparison across

simulation experiments by providing a stable definition of dependent measures. Finally, the most significant benefit of this method is the explicit decomposition of PUAs into correct behaviors and performance deviations, and further decomposition of performance deviations into innovative acts, competence errors, and true operational errors.

Extensions

This work may be extended to both improve the evaluation method and to identify design implications of the PUA classification system.

Extensions to Improve the Evaluation Method: The PUA database must be extended to include desirable skill-based performance, performance of piloting goals beyond "aviate, navigate, communicate," IAs, and to include additional requirements associated with new designs and infrastructure changes. Measures associated with level-of-effort must be surreptitiously included in the evaluation and used as covariates in evaluations. Biases in database entry must be minimized. The prescribed method for using this database requires sufficient interaction with expert evaluators that it will undergo review each time it is used. Recognition of these potential biases and errors, however, demand that evaluators are provided with visibility into the database, a method for submitting corrections to it, and a system for establishing configuration control. Concerns related to the generalizability of evaluations to novel situations and designs argue for a more analytical approach to evaluating the safety of flightdeck performance. There are complementary advantages and disadvantages (*e.g.*, face validity) to analytical approaches to characterizing a flightdeck's capacity to support safe performance. This empirical work may be useful to validate more analytical approaches.

Flightdeck Design Extension: Delineating PUAs into CAs, IAs, CEEs, CEUs and OEs, increases the perspicacity of that which we measure, provides insight into what contributes to such deviations, and suggests intervention methods. Deviations deemed IAs are useful to illustrate where the PUA database and training may be extended. Further, IAs indicate a flightdeck design that supports the pilot by providing sufficient information, tools, and flexibility of control authority to perform adaptively. Competency Errors inform us where either an element of the normative model has not been conveyed to the pilot through training or the pilot has not retained this information. CEs indicate aspects of piloting behavior that may require improved training or more frequent remedial training. Items that are covered in training, yet still emerge as CEs may suggest "training resistant behaviors;" despite inclusion in training, pilots fail to execute these behaviors in practice and require design support. Unexpressed competency errors indicate compensatory aspects of flightdeck design. Finally, true OEs can strictly be defined as such, and indicate where the competencies of the pilot (*e.g.*, resource limitations, incorrect calculations, *etc.*) and machine (*e.g.*, misleading or insufficient information, maladaptive automation, *etc.*) system fail.

ACKNOWLEDGMENTS

The author gratefully acknowledges the helpful reviews by Mr. Paul Schutte and Dr. Kelli Willshire and useful discussions with Mr. Hayes Press, Dr. Kurt Joseph, and Dr. Amy Pritchett.

REFERENCES

Beach, L.R., Chi, M., Klein, G., Smith, P., and Vicente, K. (1997) Naturalistic decision making and related research lines. In C.E. Zsambok & G. Klein (Eds.) *Naturalistic Decision Making.* Lawrence Erlbaum Associates: Mahway, NJ.

Chompsky, N. (1957) *Syntactic Structures.* The Hague: Mouton Publishers.

Hollnagel, E. (1991) The phenotype of erroneous actions. In G. R. S. Wier and J. L. Alty (eds) *Human-Computer Interaction and Complex Systems.* Academic Press: N.Y.

Joseph, K. (January, 1999) Personal Communication.

Joseph, K. M. and Uhlarik, J. (1997) Using the competence performance distinction to identify and assess situation awareness in a simulated IFR environment. *International Symposium of Aviation Psychology*, pp.1429-1435. Columbus, Ohio: Ohio State University.

Pachella, R. G. (1974). The interpretation of reaction time in information processing research. In B. Kantowitz (Ed.), *Human information processing: Tutorials in performance and cognition* (pp. 41-82). New York: Halstead Press.

Press, H. (1998) Lessons learned in creating the operational knowledge elements database. *Lockheed-Martin CN-NAS-1-96014.*

Rasmussen, J., Pedersen, O., Mancini, G., Carnino, A., Griffon, M., and Gagnolet, P. (1981) Classification system for reporting events involving human malfunctions. *Technical Report Risø-M-2240, SIN-DOC(81)14,* Risø National Laboratory, Roskilde, Denmark.

Reason, J. (1990) *Human Error*. Cambridge University Press:NY.

Rogers, W.H., Schutte, P.C., & Latorella, K.A. (1996). Fault management in aviation systems. In R. Parasuraman & M. Mouloua (Eds.), *Automation and Human performance: Theory and Applications.* (pp. 281-317). Mahwah, NJ: Erlbaum Assoc.

Smith, K. and Hancock, P.A. (1995) Situation awareness is adaptive, externally-directed consciousness. *Human Factors, 37(1),* 137-148.

Yeh, Y.Y. and Wickens, C.D. (1988) Dissociation of performance and subjective measures of workload. *Human Factors, 30,* 111-120.